

## Patent Abstracts of Japan

PUBLICATION NUMBER : 2001167110  
PUBLICATION DATE : 22-06-01

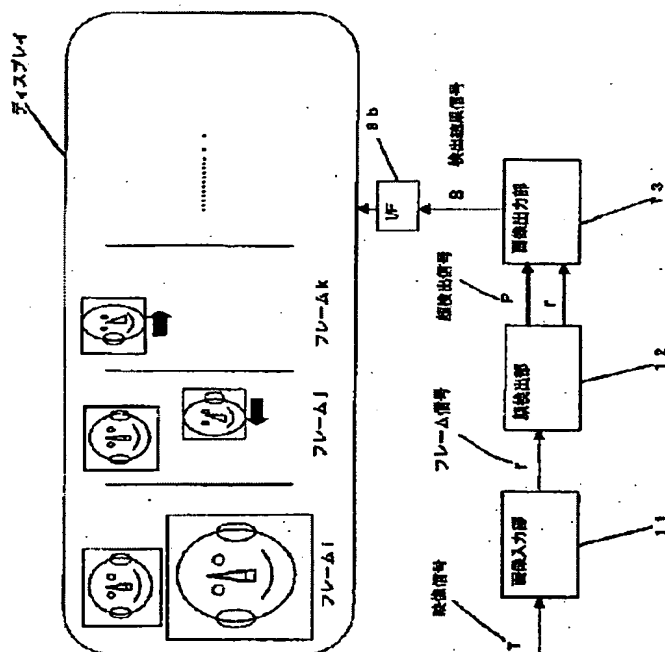
APPLICATION DATE : 08-12-99  
APPLICATION NUMBER : 11349526

APPLICANT : MATSUSHITA ELECTRIC IND CO LTD;

INVENTOR : NAGAO KENJI;

INT.CL. : G06F 17/30 H04N 5/76

TITLE : PICTURE RETRIEVING METHOD AND  
ITS DEVICE



**ABSTRACT :** PROBLEM TO BE SOLVED: To distinguish and display the faces of persons appearing in video by detecting the faces in particular from the video and identifying the detected face in addition.

**SOLUTION:** The device provided with a means for detecting a face from the video and a means for identifying the detected face detects a frame including the face from the video, extracts a face picture from the frame, and groups the faces of the same appearing person from all the extracted face pictures to extract the representative face picture of each appearing person to identify the face of the appearing person in the video. Thus, the face of the person appearing in the video can be distinguished and displayed.

COPYRIGHT: (C)2001,JPO

**BEST AVAILABLE COPY**

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-167110

(P2001-167110A)

(43) 公開日 平成13年6月22日 (2001.6.22)

(51) Int.Cl.

識別記号

F I

テマコード (参考)

G 0 6 F 17/30

H 0 4 N 5/76

B 5 B 0 7 5

H 0 4 N 5/76

G 0 6 F 15/40

3 7 0 D 5 C 0 5 2

15/403

3 5 0 C

審査請求 未請求 請求項の数14 O L (全 20 頁)

(21) 出願番号 特願平11-349526

(22) 出願日 平成11年12月8日 (1999.12.8)

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 伊藤 隆太

神奈川県川崎市多摩区東三田3丁目10番1号 松下技研株式会社内

(72) 発明者 山田 伸

神奈川県川崎市多摩区東三田3丁目10番1号 松下技研株式会社内

(74) 代理人 100082692

弁理士 蔵合 正博

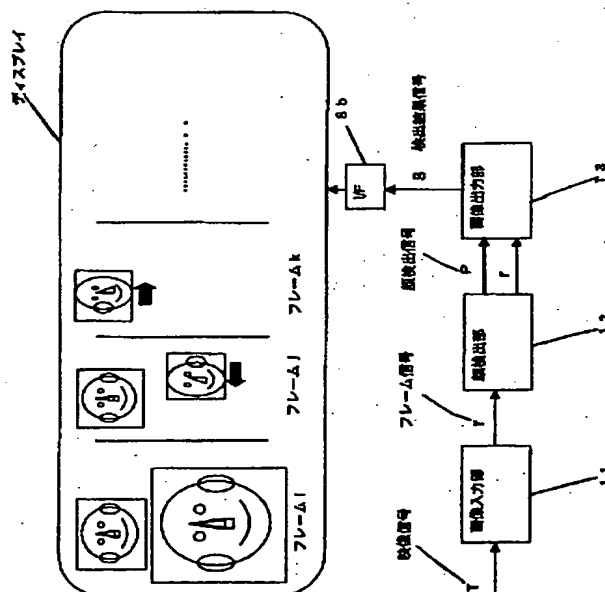
最終頁に続く

(54) 【発明の名称】 画像検索方法及びその装置

(57) 【要約】

【課題】 映像から顔に特化した検出を行い、さらに検出した顔の識別を行うことで、映像中に登場する人物の顔を区別して表示することができるようにすること。

【解決手段】 映像から顔を検出する手段と、検出した顔を識別する手段とを設け、映像中から顔が写っているフレームを検出し、前記フレームから顔画像を抽出し、抽出した全ての顔画像から同一登場人物の顔をグループ化し、登場人物別にその代表顔画像を抽出するようにし、映像中の登場人物の顔を識別する。これにより、映像中に登場する人物の顔を区別して表示することができる。



## 【特許請求の範囲】

【請求項1】 映像中から顔が写っているフレームを検出し、前記フレームから顔画像を抽出し、抽出した全ての顔画像から同一登場人物の顔をグループ化し、登場人物別にその代表顔画像を抽出することを特徴とする画像検索方法。

【請求項2】 映像から顔が写っているフレームの検出は、映像からシーンの切れ目を検出し、各シーンの代表画像として顔が含まれているフレームを検出することを特徴とする請求項1記載の画像検索方法。

【請求項3】 映像から顔が写っているフレームの検出は、写っている顔の数（人数）、顔のサイズ、顔の向き、性別、顔の表情、（年齢推定、顔の人種判定、）眼鏡の有無、またはひげの有無のうちの少なくとも1つ以上を所定の条件として検出することを特徴とする請求項1または2記載の画像検索方法。

【請求項4】 映像中の登場人物の顔と検索者が指定した顔画像との類似度を計算し、所定の類似度以上の顔が写ったフレームを検出することを特徴とする画像検索方法。

【請求項5】 映像中の登場人物の顔は、映像からシーンの切れ目を検出し、各シーンの代表画像として顔が含まれているフレームを選択することを特徴とする請求項4記載の画像検索方法。

【請求項6】 映像から人物が写っているフレームを検出し、所定の条件を満たす人物の姿勢または服装の写ったフレームを抽出することを特徴とする画像検索方法。

【請求項7】 映像から人物が写っているフレームの検出は、映像からシーンの切れ目を検出し、各シーンの代表画像として顔が含まれているフレームを検出することを特徴とする請求項6記載の画像検索方法。

【請求項8】 検索者が指定する顔画像は、登場人物リストにより予め登録した顔データベースから少なくとも1つ以上の顔画像を指定することを特徴とする請求項4記載の画像検索方法。

【請求項9】 登場人物リストは、予め検索者が作成するか、または番組表から生成することを特徴とする請求項8記載の画像検索方法。

【請求項10】 検索者が指定する顔画像は、予め映像中の登場人物の顔画像を指定登録することを特徴とする請求項4記載の画像検索方法。

【請求項11】 映像中から顔が写っているフレームを検出し、前記フレームから少なくとも1つ以上の顔画像を抽出する顔検出部と、前記抽出した全ての顔画像から同一登場人物の顔をグループ化し、登場人物別にその代表顔画像を抽出する登場人物識別部とを有することを特徴とする画像検索装置。

【請求項12】 映像信号を入力としてフレーム画像を時系列に監視することによりシーンの切り変わりを検出するシーンチェンジ検出部と、前記シーンチェンジ検出

部から各シーンの代表画像から顔が写っているフレームを検出し、前記フレームから少なくとも1つ以上の顔画像を抽出する顔検出部と、前記抽出した全ての顔画像から同一登場人物の顔をグループ化し、登場人物別にその代表顔画像を抽出する登場人物識別部とを有することを特徴とする画像検索装置。

【請求項13】 検索者が登場人物リストにより予め登録した顔データベースから少なくとも1つ以上の顔画像を指定する登場人物指定部と、映像信号を入力として受けてフレーム単位で出力する映像入力部と、映像入力部から出力されたフレーム画像から顔領域を検出する顔検出部と、前記登場人物指定部から出力された顔画像と前記顔検出部から顔領域で検出された顔画像とから登場人物に該当するか否かを判定する登場人物識別部と、前記登場人物識別部から判定された識別結果を表示・記録する画像出力部とを有することを特徴とする画像検索装置。

【請求項14】 検索対象の画像を読み込む顔画像読み込み部と、映像信号を映像信号を入力として受けてフレーム単位で出力する映像入力部と、映像入力部から出力されたフレーム画像から顔領域を検出する顔検出部と、前記検索対象画像と前記顔検出部から出力された顔領域で検出した顔画像が検索対象あるか否かを判定する照合部と、前記照合結果を表示・記憶する画像出力部とを有することを特徴とする画像検索装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、映像編集装置や画像検索装置において、動画像中における特定物が写っている画像を探索する方法及び装置に関する。

## 【0002】

【従来の技術】従来の画像検索装置として、特開平6-223179号公報に記載されたものが知られている。上記出願は、カラー動画像から特定物（検索対象）が写っているフレームを検出する画像検索装置である。図1はその構成を示すブロック図である。図1において、1は特定対象物を判定する中央処理装置としてのコンピュータ、2はコンピュータ1の出力画面を表示するCRT等のディスプレイ、3は光ディスク等の動画像再生装置、4はアナログ信号をデジタル信号に変換するA/D変換器、5は動画像再生装置3とコンピュータ1間の制御信号を接続する制御線、6はハードディスクなどからなる外部記憶装置、7はマウスなどの入力装置、8a～8eは、コンピュータ1と周辺装置との接続を行うインタフェース、9はコンピュータにおける演算処理を行なうCPU、10はCPU9から直接アクセスするメモリである。

【0003】図12は、上記従来の画像検索装置の動作を示すフローチャートである。以下、図12のフローチャートに従って動作を説明する。まず、検索者が検索したい画像を選択し、装置に入力装置7などから入力する

と、対象物を含む画像が1枚指定される(ステップ2001)。この装置は、入力された画像について、類似した色の領域に分割する(ステップ2002)。分割された各部分領域について色のヒストグラムを生成し(ステップ2004)、度数の高い色を順にN個選択し、表1のリストCG(r)を作成する(ステップ2005)。ただし、CG(r)は、r番目の部分領域のリストを示す。さらに各部分領域について、リストCG(r)と、当該部分領域と画像中の位置関係で隣接する部分領域のリストCG(r)との対応関係を表2に示すリストRCGPを作成する(ステップ2007)。

【0004】次に探索の動作について説明する。図13は、探索動作を示すフローチャートである。まず、判定対象のフレームを複数のセルC(x, y)分割し(201)、各セルについて色ヒストグラムを生成し(ステップ203)、ヒストグラムの度数が設定閾値より大きい色をリストCC(x, y)に登録する(ステップ205、206)。次に、各セルC(x, y)について(ステップ208)、RCGPでの色対を構成する色群CGうちの一方の色群に属するいずれかの色が、リストCC(x, y)に含まれており(ステップ209)、他方の色群に属するいずれかの色が、セルC(x, y)自身、あるいはその隣接する8セルのうちのいずれかのセルのリストCCに含まれている場合は、セルC(x, y)を有効セルとして抽出し(ステップ210)、1フレームにおける有効セルの総数を求める(ステップ211)。1フレームにおける有効セルの数が閾値以上であれば(ステップ212)、その色群対CGを有効色群対としてカウントする(ステップ213)。そして、カウントした有効色群対の総数が設定閾値以上ならば、対象物(検索対象)フレーム内に存在すると判定する(ステップ215)。閾値以下ならば、対象物が存在しないものとして、判断する(ステップ216)というものである。

【0005】この画像検索装置においては、映像編集者または画像検索者の負担を軽くし、効率よく検索対象の画像を探し出す技術が要求されている。しかるに、上記従来の技術では、検索したい画像の色群化と色群間での位置関係の情報に基いて検索するので、例えば制服等、服装が同一の場合、実際は検索したい人物と異なる人物あるのかかわらず、検索対象に人物であると判定しまい、誤検出の増加を招いてしまうという課題を有している。

【0006】本発明は、上記従来の問題点に鑑みてなされたもので、映像から顔に特化した検出を行い、さらに検出した顔の識別を行うことで、上記課題を解決するものである。

【0007】

【課題を解決するための手段】この課題を解決するために、本発明は、映像から顔を検出する手段と、検出した

顔を識別する手段とを設けることにより、誤検出が少なく、同じ服装であっても異なる人物であると識別し得るようにしたことを要旨とするものである。

【0008】このような態様を有する発明として、本発明請求項1に記載の発明は、画像検索方法として、映像中から顔が写っているフレームを検出し、前記フレームから顔画像を抽出し、抽出した全ての顔画像から同一登場人物の顔をグループ化し、登場人物別にその代表顔画像を抽出するようにしたものであり、映像中の登場人物の顔を識別することを特徴とする画像検索装置としたものであり、映像中に登場する人物の顔を区別して表示することができるという作用を有する。

【0009】本発明の請求項2に記載の発明は、請求項1記載の画像検索方法において、映像から顔が写っているフレームの検出は、映像からシーンの切れ目を検出し、各シーンの代表画像として顔が含まれているフレームを検出するようにしたものであり、各シーンの代表画像として顔が写っている画像を選択することができるという作用を有する。

【0010】本発明の請求項3に記載の発明は、請求項1または2記載の画像検索方法において、映像から顔が写っているフレームの検出は、写っている顔の数(人数)、顔のサイズ、顔の向き、性別、顔の表情、(年齢推定、顔の人種判定、)眼鏡の有無、またはひげの有無のうちの少なくとも1つ以上を所定の条件として検出するようにしたものであり、所定の条件を満たす顔画像を抽出することにより検索者の指定した条件に一致する顔画像が含まれた画像を抽出できるという作用を有する。

【0011】本発明の請求項4に記載の発明は、画像検索方法として、映像中の登場人物の顔と検索者が指定した顔画像との類似度を計算し、所定の類似度以上の顔が写ったフレームを検出するようにしたものであり、検索者が指定した顔が映像中のどこに記録されているかを知ることができるという作用を有する。

【0012】本発明の請求項5に記載の発明は、請求項4記載の画像検索方法において、映像からシーンの切れ目を検出し、各シーンの代表画像として顔が含まれているフレームを選択することを特徴とする画像検索装置としたものであり、各シーンの代表画像として顔が写っている画像を選択することができるという作用を有する。

【0013】本発明の請求項6に記載の発明は、画像検索方法として、映像から人物が写っているフレームを検出し、所定の条件を満たす人物の姿勢または服装の写ったフレームを抽出するようにしたものであり、特定の人物の属性がふくまれた画像を抽出することができるという作用を有する。

【0014】本発明の請求項7に記載の発明は、請求項6記載の画像検索方法において、映像から人物が写っているフレームの検出は、映像からシーンの切れ目を検出し、各シーンの代表画像として顔が含まれているフレー

ムを検出するようにしたものであり、各シーンの代表画像として顔が写っている画像を選択することができるという作用を有する。

【0015】本発明の請求項8に記載の発明は、請求項4記載の画像検索方法において、検索者が指定する顔画像は、登場人物リストにより予め登録した顔データベースから少なくとも1つ以上の顔画像を指定するようにしたものであり、人物の特定が容易に行なえるという作用を有する。

【0016】本発明の請求項9に記載の発明は、請求項8記載の画像検索方法において、登場人物リストは、予め検索者が作成するか、または番組表から生成するようにしたものであり、予め作成することにより検索者による顔画像の指定が簡単にできるとともに、番組表から生成することにより登場人物リストが簡単且つ容易に生成できるという作用を有する。

【0017】本発明の請求項10に記載の発明は、請求項4記載の画像検索方法において、検索者が指定する顔画像は、予め映像中の登場人物の顔画像を指定登録するようにしたものであり、顔画像を予め指定登録することにより検索者による顔画像の指定が簡単にできるという作用を有する。

【0018】本発明の請求項11に記載の発明は、画像検索装置として、映像中から顔が写っているフレームを検出し、前記フレームから少なくとも1つ以上の顔画像を抽出する顔検出部と、前記抽出した全ての顔画像から同一登場人物の顔をグループ化し、登場人物別にその代表顔画像を抽出する登場人物識別部とを備えたものであり、映像中に登場する人物の顔を区別して表示することができるという作用を有する。

【0019】本発明の請求項12に記載の発明は、画像検索装置として、映像信号を入力としてフレーム画像を時系列に監視することによりシーンの切り変わりを検出するシーンチェンジ検出部と、前記シーンチェンジ検出部から各シーンの代表画像から顔が写っているフレームを検出し、前記フレームから少なくとも1つ以上の顔画像を抽出する顔検出部と、前記抽出した全ての顔画像から同一登場人物の顔をグループ化し、登場人物別にその代表顔画像を抽出する登場人物識別部とを備えたものであり、各シーン毎に顔の写っているフレーム代表画像としてディスプレイに表示したり、記憶したりできるという作用を有する。

【0020】本発明の請求項13に記載の発明は、画像検索装置として、検索者が登場人物リストにより予め登録した顔データベースから少なくとも1つ以上の顔画像を指定する登場人物指定部と、映像信号を入力として受けてフレーム単位で出力する映像入力部と、映像入力部から出力されたフレーム画像から顔領域を検出する顔検出部と、前記登場人物指定部から出力された顔画像と前記顔検出部から顔領域で検出された顔画像とから登場人

物に該当するか否かを判定する登場人物識別部と、前記登場識別部から判定された識別結果を表示・記録する画像出力部とを備えたものであり、ある映像番組に登場人物を登場人物別にディスプレイに表示したり記憶したりすることができるという作用を有する。

【0021】本発明の請求項14に記載の発明は、画像検索装置として、検索対象の画像を読み込む顔画像読み込み部と、映像信号を映像信号を入力として受けてフレーム単位で出力する映像入力部と、映像入力部から出力されたフレーム画像から顔領域を検出する顔検出部と、前記検索対象画像と前記顔検出部から出力された顔領域で検出した顔画像が検索対象あるか否かを判定する照合部と、前記照合結果を表示・記憶する画像出力部とを備えたものであり、検索者が指定した顔画像に一致もしくは類似した顔画像を含むフレームをディスプレイに表示したり記憶したりすることができるという作用を有する。

【0022】

【発明の実施の形態】（実施の形態1）以下、本発明の実施の形態について、図1から図11を用いて説明する。図1乃至図3は本発明の第1の実施の形態に係る画像検索装置を説明する図である。本実施の形態では、顔の検出に関し説明したものである。具体的には、読み込んだ画像に顔が存在するか否かを判定し、顔領域が存在する場合は、画像中の顔の位置とその数、顔のサイズ、顔の向き（正面、上向き、下向き、左向き、右向き）の判別し、その結果をモニタ等に表示するものである。

【0023】図1は、顔の検出装置の全体構成図である。図1において、1は対象画像を検出するコンピュータであり、2はコンピュータ1の検出結果を表示するCRT等であり、3は光ディスク等の動画画像再生装置であり、4はアナログ信号をデジタル信号に変換するA/D変換器であり、5は動画画像再生装置3とコンピュータ1間の制御信号を接続する制御線、6はハードディスクなどからなる外部記憶装置であり、7はマウス、キーボードなどの入力装置、8a~8eは、コンピュータ1と周辺装置との接続を行うインタフェース、9はコンピュータのCPU、10はCPU9から直接アクセスするメモリである。

【0024】動画画像再生装置3から出力される映像信号は、逐次、A/D変換器4によってデジタル画像に変換され、コンピュータ1に送られる。デジタル画像は、コンピュータ1では、インタフェース8cを介してメモリ10に入り、メモリ10に格納されたプログラムに従って、CPU9により処理される。

【0025】図2は検出方法の流れを示したフローチャートである。まず、フレーム画像rがコンピュータ1に読み込まれると（ステップ101）、予めプログラムによって記述された関数にfに従って、顔領域が検出される（ステップ102）。関数fの出力Pはフレーム上で

検出した全ての顔座標が記録された行列である。例えば、行列Pのi列目のベクトルは、フレームrでi番目に検出された顔を格納しており、行列Pの(i,1)成分は、例えば顔領域が長方形の領域で切り出されるとすれば、顔領域の左上のx座標、(i,2)成分は顔領域の左上のy座標、(i,3)成分は顔領域の右下のx座標、(i,4)成分は、顔領域の右下のy座標を格納している。また、顔領域が1つも見つからなかった場合は、関数fは例えば-1を返すようにしておく。行列Pの出力が-1でなければ、ディスプレイ2に検出結果が出力される(ステップ104)。

【0026】ディスプレイには検出した顔領域のみ表示されても良いし、顔領域に特定マークが付されてフレーム全体が表示されても良い。なお、行列Pの列サイズが検出した人数を表しているため、検出人数を併せてディスプレイに表示することもできる。

【0027】具体的な関数fについては、例えばカラー画像の場合は、肌色検出により実現できる。入力された画像の各画素の(R, G, B)値を色空間(例えば、Yuv空間)にプロットし、予め定義した肌色空間に含まれた画素のみを選択する。さらに、前述の該当画素のみで形成される画像領域のうち、予め設定した閾値以上の面積を持った領域で且つ、その領域の形状が予め定めた形状(例えば楕円)と同一または類似する領域のみを選択することにより、顔検出を実現する。また、モノクロ画像とカラー画像のどちらにも対応できる方法として、テンプレートマッチングによる方法も考えられる。予め1枚または複数の顔画像を標準パターンとして記憶しておく。また、フレーム画像に対しては、切り出し用ウィンドウの位置を移動させながらフレーム上の一部の領域を順次切り出していき、切り出し画像と上述のテンプレートとの類似度を計算する。例えば、切り出し画像とテンプレート画像の相関値が設定閾値以上の場合には、その切り出し領域は、顔画像であると判定する。

【0028】なお、異なるサイズのテンプレートを用意しておくことにより、顔のサイズの異なる顔領域も検出することが可能である。また、顔の向きの異なるテンプレートを用意しておくことにより、顔の向きの異なる顔領域も検出することができる。例えば、上向き・下向き・左向き・右向きの4種類の顔画像のテンプレートを用意しておき、各切り出し画像に対して上記4種類のテンプレートとの類似度を求め、この4つの類似度の最大値が予め定めた閾値以上であれば、該当するテンプレートから顔の向きを特定することができる。

【0029】別の関数fの実現方法として、切り出された領域が、顔か非顔であるかの識別関数を予め作成しておく方法も考えられる。例えばニューラルネットワークを用いて学習により、顔と非顔の識別関数を作成しておくというものである。なお、顔の向きの異なる画像に予め異なるラベルをふることにより形成したカテゴリを識

別するようにニューラルネットワークを学習させ、顔の向きを検出することもできる。

【0030】図3は、顔検出の処理をブロック図で示したものである。図3において、11は、画像入力部であり、12は顔検出部であり、13は画像出力部である。画像入力部は、A/D変換器4から出力される映像信号Tを入力として、1フレーム分の画像をフレーム信号rとして出力する。顔検出部12は、画像入力部から出力されたフレーム信号rを入力として、1フレーム内に存在する顔の位置の座標、顔の向きの情報を出力する顔検出信号Pと、フレーム信号rを出力する。画像出力部13は、顔検出部12から出力される顔検出信号Pとフレーム信号rとから、顔領域のみの画像を検出結果信号Sとして出力する。検出結果信号Sは、コンピュータ内のインタフェース8bを介してディスプレイ2に表示される。図2のフレームには、例えば各フレーム毎に切り出された顔画像が表示される。図2のように、様々な顔の大きさと向きに対応が可能である。例えば顔の向きが矢印等でディスプレイ上に表示されても良い。なお、検出信号Sは、顔領域に特定のマークを付されたフレーム全体の画像でも構わない。また、顔検出と同様に手法により、ひげや眼鏡の検出等もできる。

【0031】(実施の形態2)図4は、本発明の第2の実施の形態に係る画像検索装置の動作を説明するフローチャートである。この実施の形態に係る画像検索装置は上記第1の実施の形態に係る装置構成と同じ構成を有する。第2の実施の形態では、映像のシーンチェンジ機能と実施の形態1で記述した顔検出機能とを組み合わせ、或るシーンでの代表画像を選択する際に、人の顔が写っている画像を代表画像とするものである。以下図4に示したフローチャートで説明する。

【0032】図4において、まず現在のシーンの代表画像がすでに記憶されているか否かのフラグFlgを初期化する(ステップ107)。次に映像のr番目のフレームrを読み込み、シーンが変化したか否かを判断する(ステップ109)。シーンが変化したと判断された場合は、予め記憶している前シーンの代表画像RTをディスプレイに表示する(ステップ110)とともに、フラグFlgを0にリセットする(ステップ111)。ただし、映像入力の開始直後は、前シーンの代表画像RTは存在しないので、ディスプレイには何も表示されない。

【0033】次に、フラグFlgを評価し(ステップ112)、Flgが0の場合は現在読み込んでいるフレームrを代表画像RTとして記録して(ステップ113)、フラグFlgを1にセットする(ステップ114)。この処理は、該当シーン内に1度も人の顔が含まれていない場合を考慮したもので、予め、シーンチェンジが判断された直後のフレームを代表画像として記憶しておくものである。なお、シーンチェンジが判断されてから所定フレーム経過後のフレームを予め代表画像として記憶するも

のであっても良い。

【0034】フラグが1にセットされている場合は、すでに以前のフレームが代表画像を記録していることになるので、そのまま顔検出を行う。また、ステップ113、114で最初の代表画像が取り込まれた後も、実施の形態1に記述した手法で顔の検出を行う（ステップ115）。

【0035】検出の結果、関数 $f$ の出力 $P$ が-1の場合（ステップ116）、つまり顔が検出されない場合は、次のフレームを読み込む。 $P$ が-1でない場合、つまり現在読み込んでいるフレーム $r$ に顔が検出された場合は、代表画像を更新する（ステップ117）。

【0036】以上の処理手順により、あるシーンに顔領域が存在している場合は、顔が写っているフレームを代表画像とすることができる。なお、実施の形態1で記述した手法は、顔の向き、サイズ、数も判定することができるので、例えば正面顔のフレームを代表画像にするとか、顔のサイズがより大きいフレームを代表画像とするとか、顔の数がより多いフレームを代表画像として選ぶこともできる。

【0037】図5は、上記の処理手順をブロック図で表したものである。図5において、11は、画像入力部であり、実施の形態1で記述したものと同一である。14は、シーンチェンジ検出部であり、12は顔検出部であり、実施の形態1で記述したものと同一である。13は画像出力部であり、実施の形態1で記述したものと同一である。

【0038】画像入力部11は、A/D変換器4から出力される映像信号 $T$ を入力として、1フレーム分の画像をフレーム信号 $r$ として出力する。シーンチェンジ検出部14では、映像の不連続性からシーンの切り変わりを判定し、シーン切り換え信号 $C$ と、フレーム信号 $r$ を出力する。顔検出部12は、シーンチェンジ検出部14から出力されたシーン切り換え信号 $C$ とフレーム信号 $r$ を入力として、1フレーム内に存在する顔の位置の座標、顔の向きの情報を出力する顔検出信号 $P$ と、フレーム信号 $r$ を出力する。画像出力部13は、顔検出部12から出力される顔検出信号 $P$ とフレーム信号 $r$ とから、顔の写っているフレームを検出結果信号 $S$ として出力する。検出結果信号 $S$ は、コンピュータ内のインタフェース8bを介してディスプレイ2に表示される。ディスプレイ2には、例えばシーン毎の代表画像が表示され、そのシーン内に顔画像が含まれている場合は、顔が含まれている画像を代表画像にすることができる。

【0039】なお、シーンチェンジ検出部の実現は、例えば文献「ビデオインデックス作成編集技術」、(Matsushita Technical Journal Vol.44 No.5)、山田伸 他に記載された公知技術を利用することにより実現できる。

【0040】（実施の形態3）図6は、本発明の第3の実施の形態に係る画像検索装置の動作を説明するフロー

チャートである。この実施の形態に係る画像検索装置は上記第1の実施の形態に係る装置構成と同じ構成を有する。本実施の形態は、読み込んだフレームについて顔を検出し、その検出した顔から性別の識別の実施について説明したものである。

【0041】図6のフローチャートを用いて説明する。まず、フレーム画像 $r$ がコンピュータ1に読み込まれると（ステップ119）、予めプログラムによって記述された関数に $f$ に従って、顔領域が検出される（ステップ120）。関数 $f$ は、実施の形態1で説明したものと同一である。関数 $f$ の出力 $P$ が-1でない場合、すなわち、顔領域が検出された場合は、検出領域の顔が男性か、女性かの識別を行う（ステップ122）。関数 $g$ は男女識別関数である。関数 $g$ はフレーム上の顔領域の座標 $P$ と、フレーム $r$ を引数として、男女識別し、識別結果を出力 $Q$ に格納する。出力 $Q$ は例えば行列となっており、行列 $Q$ の $i$ 列目のベクトルは、 $i$ 番目に検出した顔について記録されている。行列 $Q$ の $(i, 1)$ 成分は $i$ 番目に検出した顔領域の左上の $x$ 座標、 $(i, 2)$ 成分は $y$ 座標、 $(i, 3)$ 成分は顔の右下の $x$ 座標、 $(i, 4)$ 成分は顔の右下の $y$ 座標を、 $(i, 5)$ 成分は、男性の場合は+1、女性の場合は-1が格納される。

【0042】関数 $g$ の実現方法としては、予め性別が既知の顔画像を多数収集し、この収集した顔画像を元に判別分析等の統計的手法により実現することができる。また例えば、ニューラルネットワークの学習により、識別関数を実現することもできる。最後に、ディスプレイ2に検出結果が出力される（ステップ123）。

【0043】したがって、本実施の形態では、顔検出したのち、さらに検出した顔が男性か女性を表示することができる。また、予め映像編集者、映像検索者が特定の性別のみを検索したい旨を入力装置7介して入力していればその特定性別のみの顔画像が含まれたフレームを検出することもできる。

【0044】図7は、上記の処理をブロック図で表示したものである。図7において、11は、画像入力部であり、12は顔検出部であり、18は男女識別部であり、13は画像出力部である。画像入力部11は、A/D変換器4から出力される映像信号 $T$ を入力として、1フレーム分の画像をフレーム信号 $r$ として出力する。顔検出部12は、画像入力部から出力されたフレーム信号 $r$ を入力として、1フレーム内に存在する顔の位置の座標、顔の向き、サイズ等の情報を出力する顔検出信号 $P$ と、フレーム信号 $r$ を出力する。男女識別部では、顔検出部12から出力された顔検出信号 $P$ と、フレーム信号 $r$ とを入力とし、 $P$ に記録された顔領域の画像に基いて男女識別を行い、識別結果と顔検出信号 $P$ との結果を併せて男女識別信号 $Q$ として出力する。フレーム信号 $r$ も出力する。画像出力部13は、男女識別部15から出力された男女識別信号 $Q$ とフレーム信号 $r$ とから、顔領域のみ

を切り出した後、各顔画像に男性、女性の識別マークを付加した画像を検出結果信号Sとして出力する。検出結果信号Sは、コンピュータ内のインタフェース8bを介してディスプレイ2に表示される。図7のディスプレイには、例えば各フレーム毎に切り出された顔画像と各顔画像の性別マークが付されて表示される。

【0045】なお、ディスプレイへの表示のやり方は上述したものに限ったものではなく、顔領域に特定のマークを付された後、フレーム全体がディスプレイに表示されるものでもよい。また、実施の形態2に記述したシーンチェンジ検出部14を図7の画像入力部11の直後に挿入することにより、各シーンの代表画像は、特性性別の写っている顔にするなども可能である。

【0046】さらに、図8のブロック図のように、例えば年齢識別部16、表情識別部17、人種識別部18とを加えることにより、より細かい検索、編集が可能となる。年齢識別部16は、顔検出部12から出力された顔検出信号Pとフレーム信号rとから、検出された顔から年齢を推定し、年齢識別信号yを出力する。年齢の識別は、例えば20代等の年代を出力する。年齢識別部16の実現方法としては、男女識別部を実現する手法と同様の方法で実現できる。

【0047】予め年代が既知の顔画像を多数収集し、この収集した顔画像を元に判別分析等の統計的手法により実現することができる。また例えば、ニューラルネットワークの学習により、識別関数を実現することもできる。この表情識別部17は、顔検出部12から出力された顔検出信号Pとフレーム信号rとを入力として、検出された顔から表情を推定し、表情識別信号Hを出力する。表情の識別は、例えば、「笑う」、「泣く」、「怒る」等の表情別に付けられたラベルが出力される。表情識別部17の実現方法としては、男女識別部を実現する手法と同様の方法で実現できる。

【0048】予め表情が既知の顔画像を多数収集し、この収集した顔画像を元に判別分析等の統計的手法により実現することができる。また例えば、ニューラルネットワークの学習により、識別関数を実現することもできる。

【0049】人種識別部18は、顔検出部12から出力された顔検出信号Pとフレーム信号rとを入力として、検出された顔から人種を推定し、人種識別信号Lを出力する。人種の識別は、例えば、「黄色人種」、「白人」、「黒人」等の人種別に付けられたラベルが出力される。表情識別部17の実現方法としては、男女識別部を実現する手法と同様の方法で実現できる。

【0050】予め人種が既知の顔画像を多数収集し、この収集した顔画像を元に判別分析等の統計的手法により実現することができる。また例えば、ニューラルネットワークの学習により、識別関数を実現することもできる。

【0051】画像出力部13は、男女識別部15から出力された男女識別信号Q、フレーム信号r、年齢識別部から出力された年齢識別信号y、表情識別部17から出力された表情識別信号Hと、人種識別部から出力された人種識別信号Lとを入力として受けて、各顔画像に性別、年齢、表情、人種のラベルを付し、コンピュータ内のインタフェース1/F8bを介してディスプレイ2に表示させる。

【0052】その他同様の手法により、髪型や帽子の識別等も可能である。また、切り出した顔画像の周囲の画像を切り出すことにより、検出された人物についてのその他の情報も取得できる。例えば、顔画像の下に位置する画像から、その人物の着ている洋服の色、ネクタイの色、洋服の種類等も識別できる。さらに背景差分等の手法からその人物の姿勢等も検出でき、複数の連続するフレームからその人物の姿勢変化を捉えることにより、動作を推定することもできる。

【0053】また、第2の実施の形態において記述したシーンチェンジ検出部14を図8の画像入力部11の直後に挿入することにより、各シーンの代表画像を、特定性別でかつ特定年代でかつ特定表情でかつ特定人種の写っている顔にするなども可能である。

【0054】ここで説明した実施の形態の用途としては映像編集、検索装置を想定したが、なにもそれに限った用途だけでなく、小売店における購入者分析等にも利用できる。例えばスーパーなどの小売店のレジにカメラを設置し、レジに並んだ購入者の映像から顔を検出して男女識別、年齢識別等を実施することにより、購入商品と性別、年齢との関係の情報を取得することができる。これらの情報により仕入れ品を変更するなどして、売り上げ改善に貢献することも可能な技術である。

【0055】(実施の形態4) 図9は、本発明の第4の実施の形態に係る画像検索装置の動作を説明するフローチャートである。この実施の形態においても、画像検索装置は上記第1の実施の形態に係る装置構成と同じ構成を有する。本実施の形態は、ある番組内の登場人物の識別に関して説明したものである。ここにいう登場人物識別とは、ある番組内に登場する人物を区別して表示する機能をいう。

【0056】最初に、ある映像番組内に登場する人物のリストが既知で、かつその登場人物の顔画像が画像データベース等から取り出すことができる場合を想定して以下を説明する。図9に示す処理のフローチャートにしたがって動作を説明する。まず、番組登場人物リストを読み込む(ステップ126)。このリストは各登場人物別にID番号が与えられており、このID番号と顔画像データベースの格納アドレスが記録されている。次に、各ID番号毎に登録されている顔データベースから顔画像を読み込む。顔画像は、例えば向きやサイズ、表情の異なる画像が登録されているとする。



【0057】上記のように各登場人物別の複数の顔画像から特徴量を抽出する(ステップ128)。例えば特徴量抽出方法の実現手法として、例えばKL展開がある。つまり登場人物別の顔画像毎にKL展開を実施する。

【0058】ここまでの、登場人物識別のための準備である。

【0059】次に、対象としている番組映像を読み込み(ステップ129)、各フレーム毎に実施の形態1に記述した手法により顔画像を検出する(ステップ130)。検出した顔画像について、顔識別を実施し、登場人物との対応付けを行う(ステップ131)。顔識別手法としては例えば、部分空間法等を用いることができる。最後に、ディスプレイに登場人物識別結果を表示する(ステップ132)。

【0060】図10は、上記処理のブロック図を示したものである。図10において、19は登場人物リストであり、20は登場人物の顔画像が記録されている顔データベースであり、21は顔画像取り込み部であり、22は顔画像から特徴を抽出する特徴抽出部であり、23は登場人物の顔を識別する顔識別部であり、11は映像信号を入力として受けて、フレーム毎にフレーム信号rを出力する画像入力部であり、12は顔検出部であり、13は識別結果を出力する画像出力部であり、8bはコンピュータ内のインタフェースであり、2は識別結果を表示するディスプレイである。以下、各ブロック別に動作を説明する。

【0061】画像取り込み部21は、まず入力された登場人物リストから顔データベースに記憶されている登場人物の顔画像を取り込み、該当顔画像を顔データベース信号f\_tとして出力する。特徴抽出部22は、顔画像取り込み部から出力される顔データベース信号f\_tから、特徴量を抽出する。例えば、各登場人物別にKL展開を行い求めた固有ベクトルを特徴量信号Kとして出力する。

【0062】一方、画像入力部11はA/D変換器4を介して入力された映像信号Tを取り込み、1フレームずつ、フレーム信号rとして出力する。顔検出部12の動作は、実施の形態1に記載したものと同一である。画像入力部11から出力されたフレーム信号rから、実施の形態1に記述した手法によりフレーム内の顔領域を検出し、顔検出信号Pとフレーム信号rを出力する。登場人物識別部23は、前述の特徴量信号Kと顔検出信号Pとフレーム信号rを入力として受け、まず顔検出信号Pとフレーム信号rとから顔領域を切り出し、ベクトル化する。このベクトルに対し、例えば部分空間法等を実施、切り出された顔画像が、どの登場人物に最も類似するかを求め、その最も類似度が高い登場人物のID番号、該当する顔画像のフレーム上の位置を併せて識別結果信号Resとして出力する。またフレーム信号rも同時に出力する。

【0063】画像出力部13は、登場人物識別部23から出力された識別結果信号Resとフレーム信号rとから、ディスプレイ2上に登場人物別に顔画像を表示するように、識別結果信号Sとして出力する。なお、ディスプレイへの表示の仕方はこれに限るものではなく、ディスプレイにはフレーム全体を表示させ、登場人物別に異なるマークを付して表示させることもできる。

【0064】さらに、実施の形態2に記述したシーンチェンジ検出部14を画像入力部11の直後に挿入することにより、シーン毎に登場人物の検出結果を出力することもできる。

【0065】なお、本実施例では予め映像番組内に登場する人物のリストが既知の場合を説明したが、登場人物が既知でない場合は、まず対象とする映像全体で顔検出したのち、教師なしクラスタリングを実施することにより実現できる。また、登場人物が顔データベースに登録されていない場合で、初期のシーンでの登場人物の顔がその後の映像中のどこに登場しているかを検出したい場合は、検索者や編集者が映像の初期シーンの中から登場人物を切り出して、特徴抽出部22に入力することにより実現できる。

【0066】(実施の形態5)図11は、本発明の第5の実施の形態に係る画像検索装置の構成を示すブロック図である。本実施の形態は、ある特定人物の顔が対象映像中のどこに存在するか検出することについて説明したものである。

【0067】図11において、25は検索対象の顔画像であり、21は検索対象の顔画像を取り込む顔画像取り込み部であり、22は顔画像から特徴を抽出する特徴抽出部であり、11は、画像入力部であり、12は顔検出部であり、26は、顔検出部12から検出された顔画像から特徴量を抽出する映像信号用特徴抽出部であり、24は照合部であり、13は識別結果を出力する画像出力部であり、8bはコンピュータ内のインタフェースであり、2は識別結果を表示するディスプレイである。

【0068】以下、ブロック毎にその動作を説明する。画像取り込み部21は、検索対象の顔画像25を取り込み、前処理を実行する。例えばヒストグラム平滑化等を実施し、前処理済検索画像信号TTaとして出力する。特徴抽出部22は、顔画像取り込み部21から出力される前処理済検索画像信号TTaから所定の特徴量を特徴量を抽出する。

【0069】一方、画像入力部11はA/D変換器4を介して入力された映像信号Tを取り込み、1フレームずつ、フレーム信号rとして出力する。顔検出部12の動作は、画像入力部11から出力されたフレーム信号rから、フレーム内の顔領域を検出し顔検出信号Pとフレーム信号rを出力し、出力された信号は映像信号用特徴量抽出部26において、顔画像を切り出し、前処理を実施した後、特徴量信号kとフレーム信号rを出力する。照

合部24では、特徴抽出部22と映像信号用特徴量抽出部から出力された特徴量信号を照合し、所定の類似度以上であれば、検索対象の顔であると判定し、そのフレームを画像出力部13に出力する。画像出力部13は、入力されたフレームをディスプレイ2上にフレーム番号とともに出力する。

【0070】以上の実施の形態によれば、例えば有名な俳優、政治家等の検索をすばやく行うことができ、映像編集者や検索者の負担を軽くすることができる。

【0071】

【発明の効果】以上のように本発明によれば、大量に蓄積されている映像データベース等から、顔画像を検出し、検出した顔を識別するので、制服等、服装が同一の場合であっても誤検出することなく精度良く検索したい顔画像が含まれた画像を抽出することができるという有利な効果を有している。さらに、本発明の効果として男女識別等、検索者のより細かい検索要求にも対応することができるという効果を有する。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態に係る画像検索装置の全体構成を表すブロック図

【図2】実施の形態1の動作を説明するためのフローチャート

【図3】前記実施の形態1を説明するためのブロック図

【図4】本発明の第1の実施の形態に係る画像検索装置の動作を説明するためのフローチャート

【図5】前記実施の形態2を説明するためののブロック図

【図6】本発明の第3の実施の形態に係る画像検索装置の動作を説明するためのフローチャート

【図7】本発明の実施の形態3を説明するためのブロック図

【図8】前記実施の形態3を説明するためのブロック図

【図9】本発明の第4の実施の形態に係る画像検索装置の動作を説明するためのフローチャート

【図10】前記実施の形態4を説明するためのブロック図

【図11】本発明の第5の実施の形態に係る画像検索装置の動作を説明するためのブロック図

【図12】従来の技術を説明するためのフローチャート

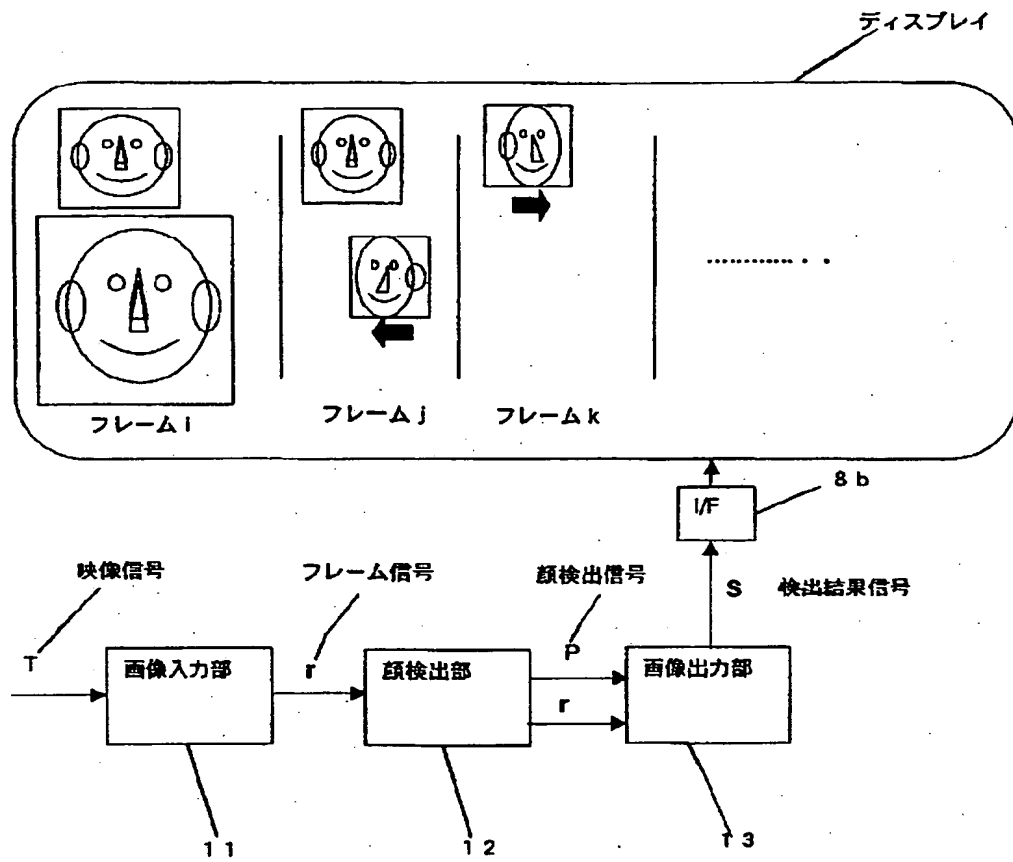
【図13】従来の技術を説明するためのフローチャート

【符号の説明】

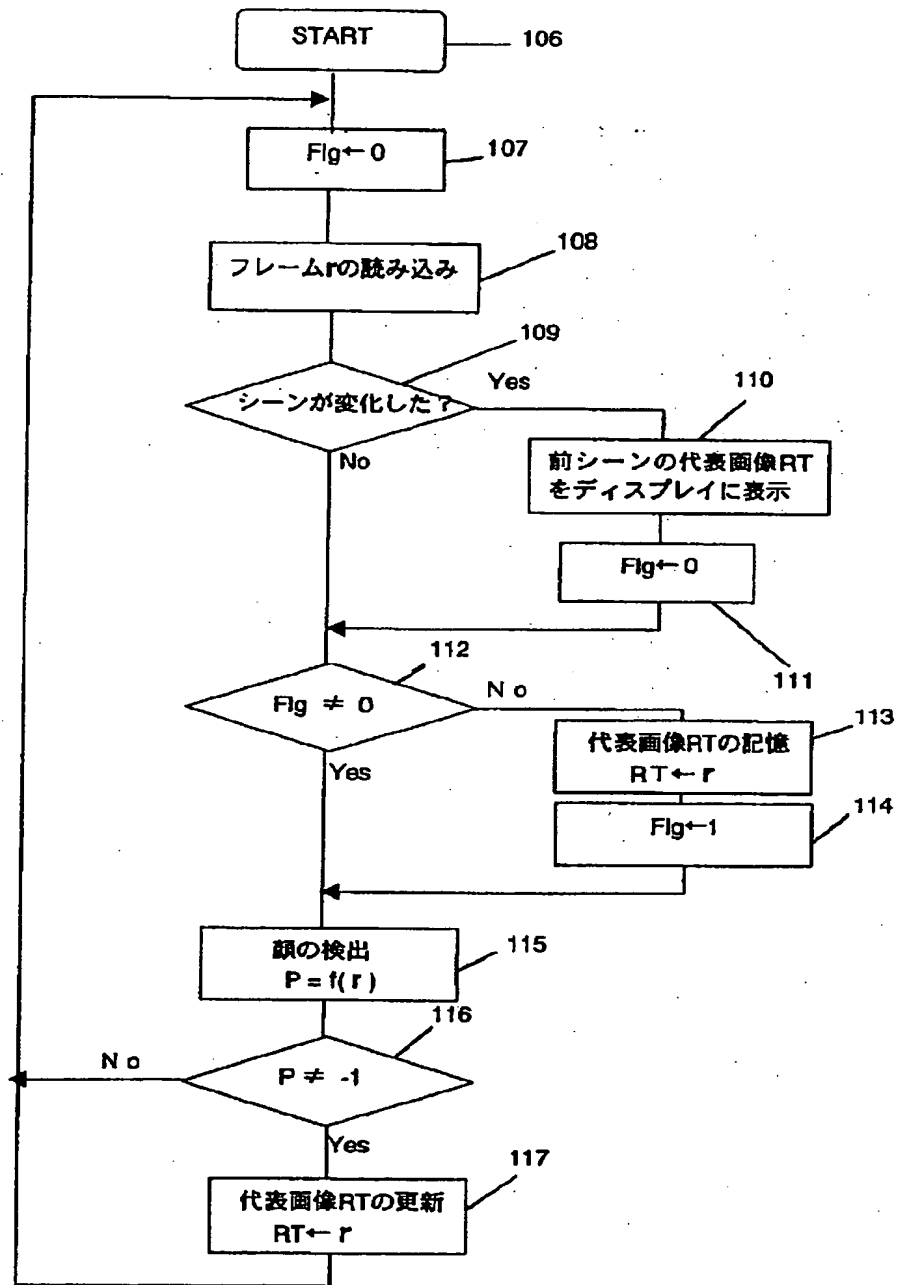
- 1 コンピュータ
- 2 ディスプレイ
- 3 動画像再生装置
- 4 A/D変換器
- 5 制御線
- 6 外部記憶装置
- 7 入力装置
- 8a~8e インタフェース
- 9 CPU
- 10 メモリ
- 11 画像入力部
- 12 顔検出部
- 13 画像出力部
- 14 シーンチェンジ検出部
- 15 男女識別部
- 16 年齢識別部
- 17 表情識別部
- 18 人種識別部
- 19 登場人物リスト
- 20 顔データベース
- 21 顔画像取り込み部
- 22 特徴抽出部
- 23 登場人物識別部
- 24 照合部
- 25 検索対象顔画像
- 26 映像信号用特徴抽出部



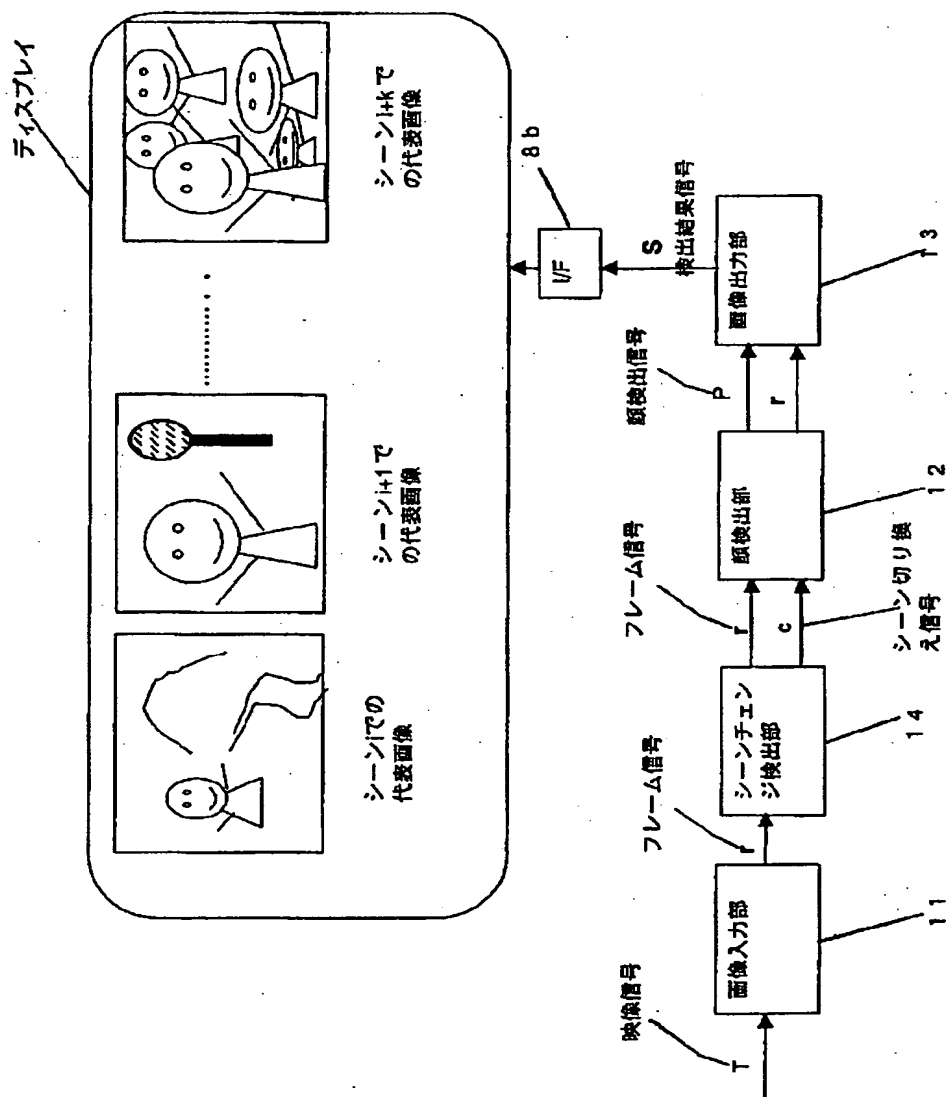
【図 3】



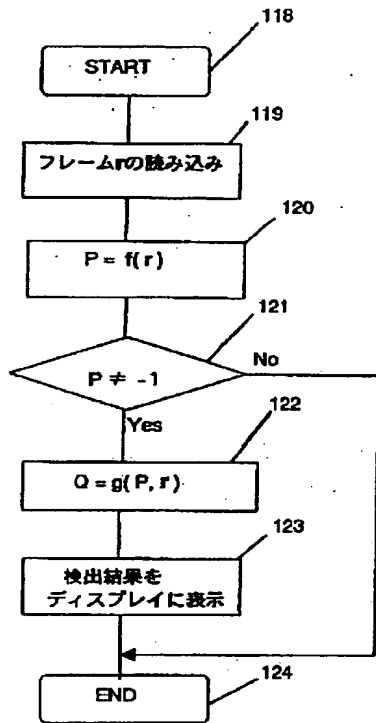
【図4】



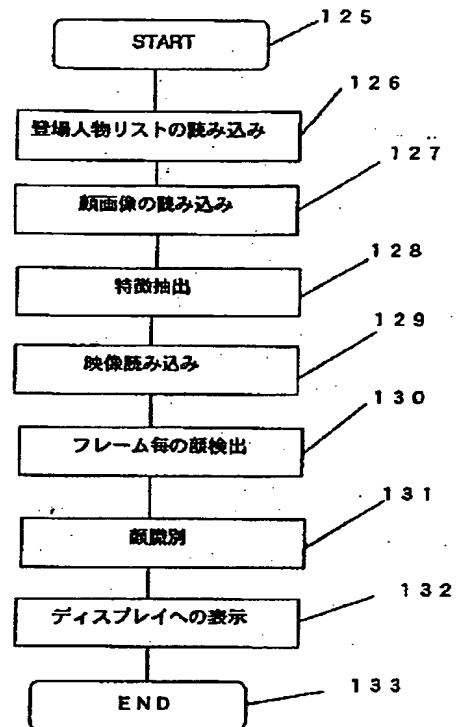
【図5】



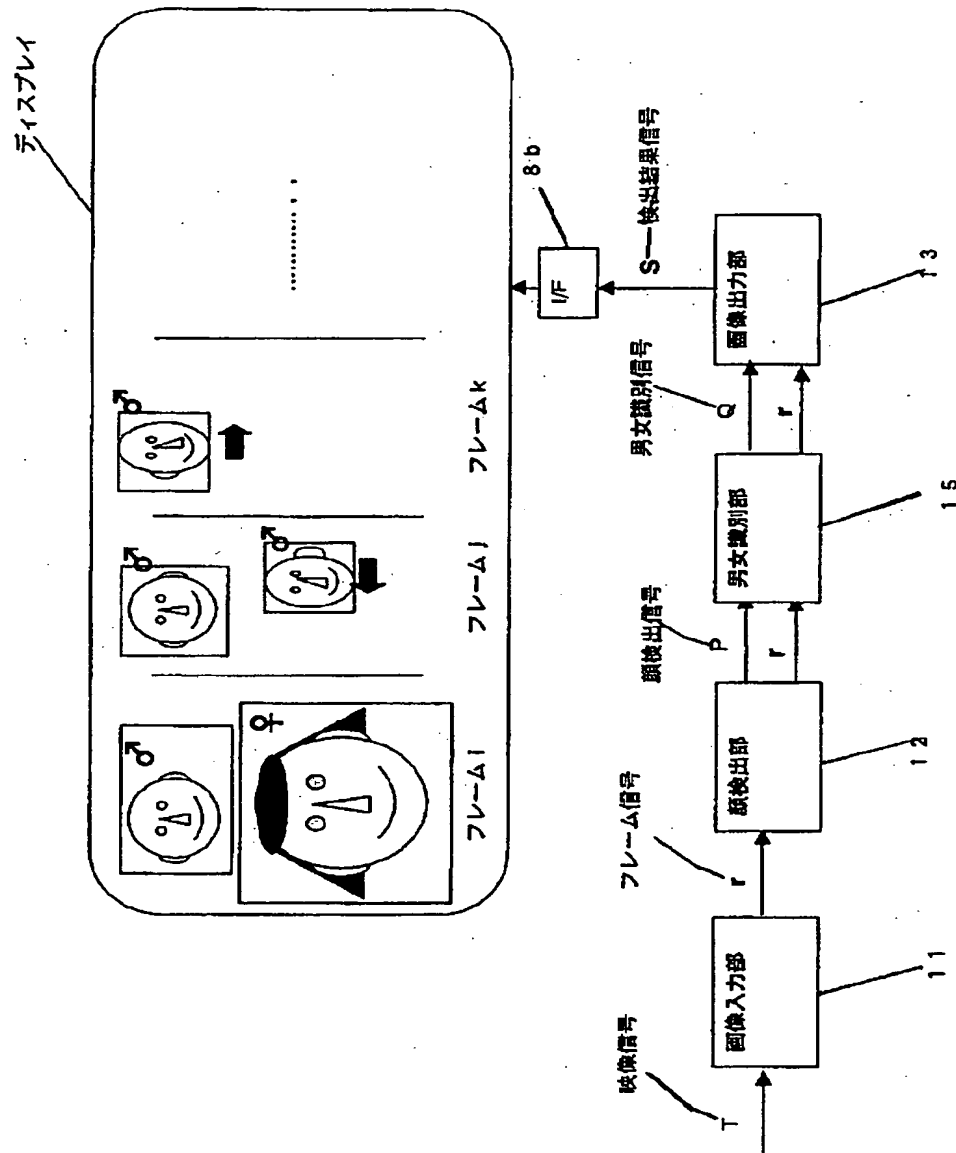
【図6】



【図9】

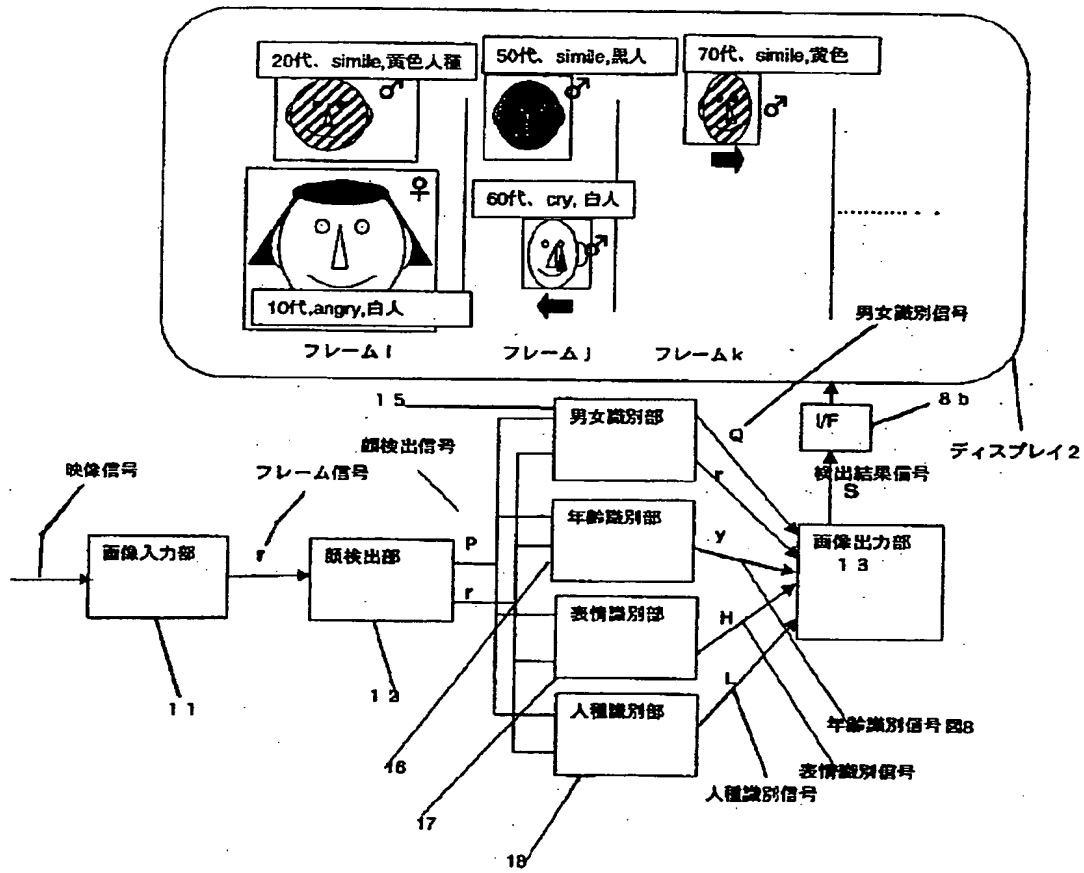


【図7】

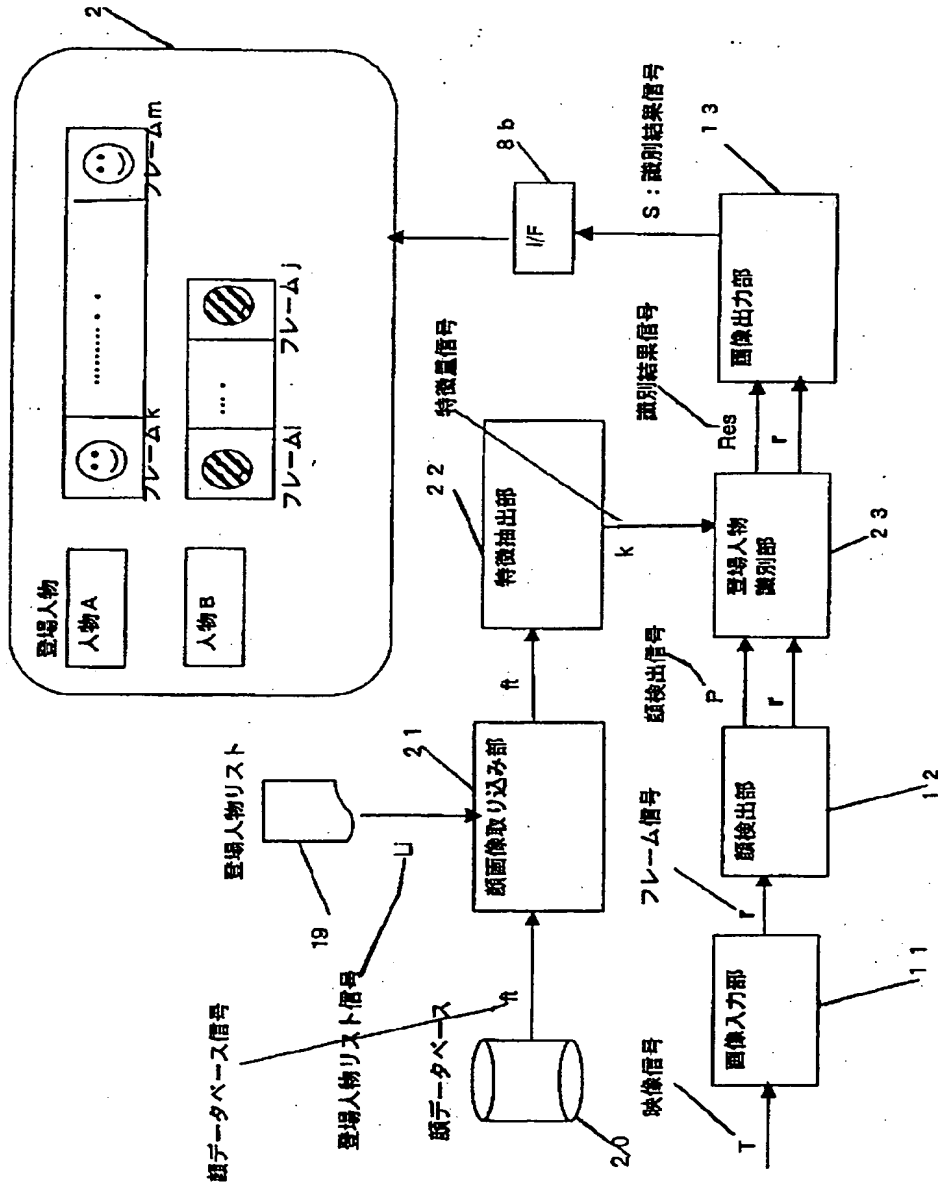




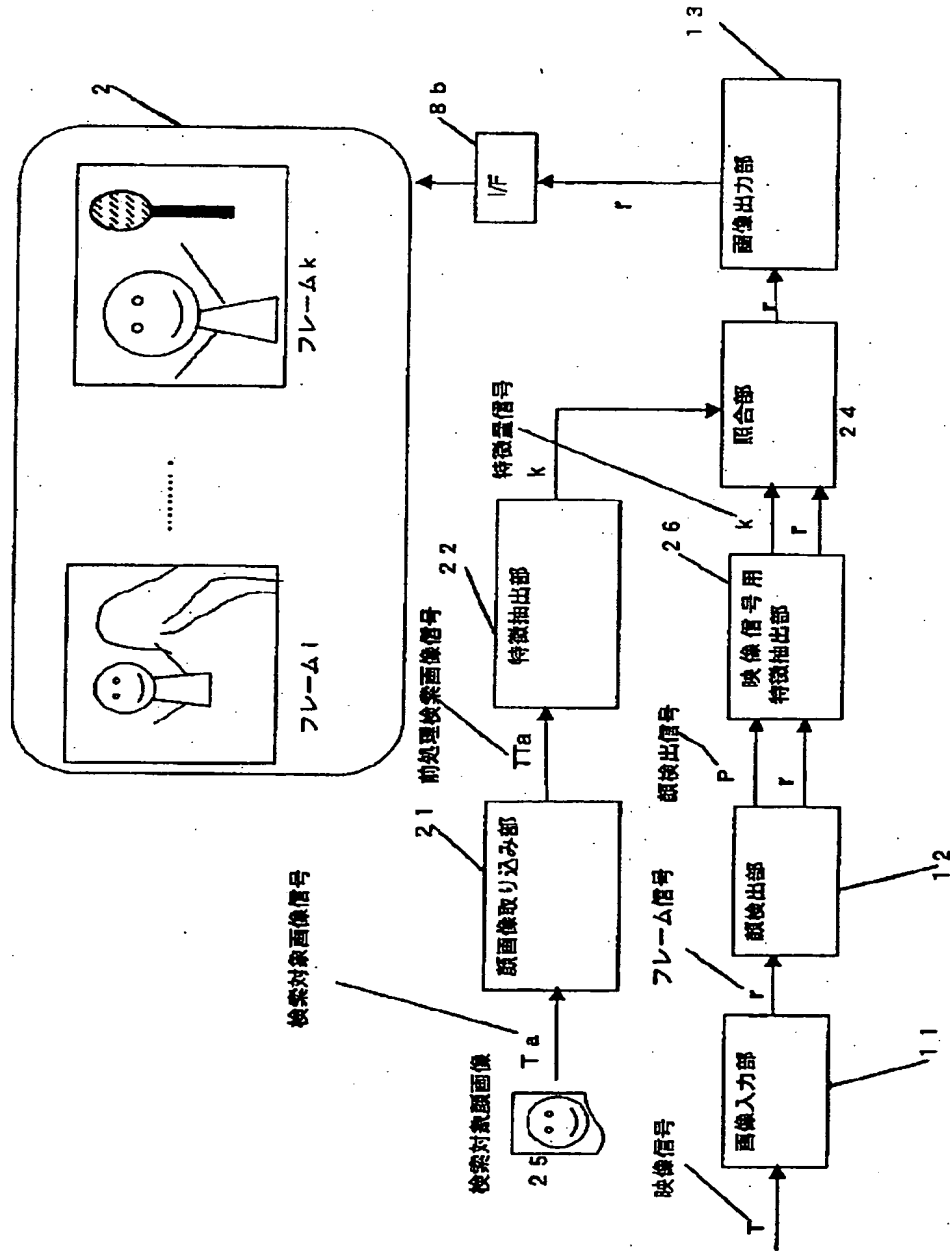
【図8】



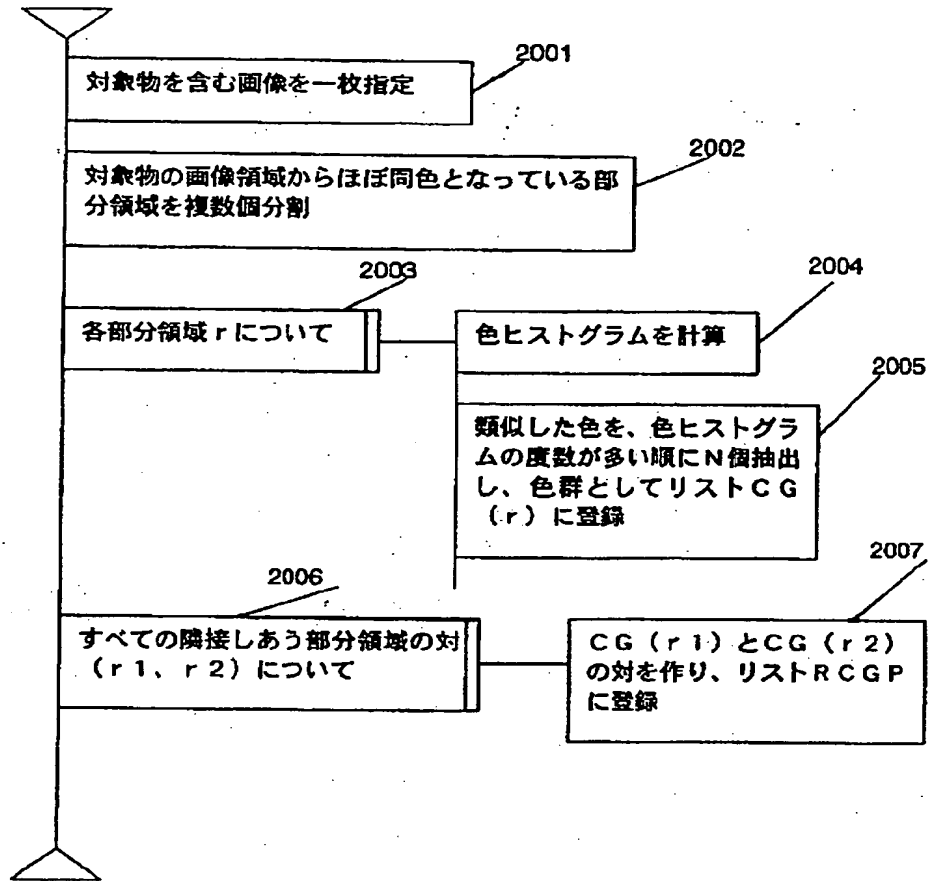
【図10】



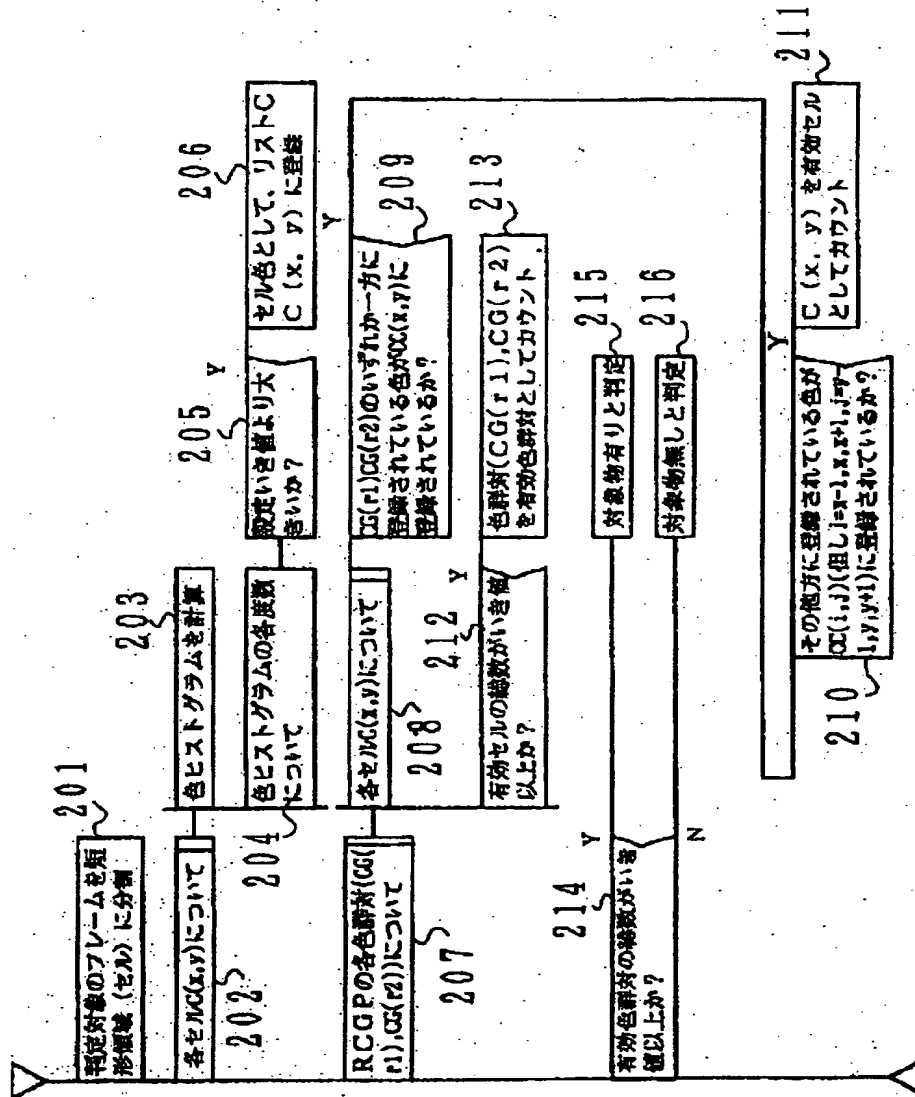
【図11】



【図12】



【図13】



フロントページの続き

(72)発明者 相馬 正宣

神奈川県川崎市多摩区東三田3丁目10番1  
号 松下技研株式会社内

(72)発明者 長尾 健司

神奈川県川崎市多摩区東三田3丁目10番1  
号 松下技研株式会社内

Fターム(参考) 5B075 ND12 NK07 PQ02 PR06 QM08  
5C052 AA01 AC08 DD04 DD10 EE02  
EE03

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☒ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☒ **FADED TEXT OR DRAWING**

☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**